REPORT

ALIGNMENT ANALYSIS OF

IDAHO LANGUAGE USAGE CONTENT STANDARDS

AND THE IDAHO STANDARDS ACHIEVEMENT TESTS

GRADES 3 – 8 AND 10

MAY 30, 2008

The findings in this study are those of the independent reviewing team and do **not** represent the opinion of the vendor Data Recognition Corporation or the State of Idaho.

Table of Contents

Executive Summary

The alignment study for language usage was held in Boise, Idaho, January 15-18, 2008. This report consists of a description of the four criteria used to judge the alignment for Grades 3 through 8 and 10 Idaho Language Usage content standards and the test items found in the Language Usage *Idaho Standards Achievement Tests*. This report also includes tables listing the results from the review process.

Eight reviewers participated in the study: four language usage expert reviewers from the State of Idaho and four national expert reviewers. A national expert facilitated the alignment process for language usage. The four State of Idaho reviewers have extensive teaching experience in the state and expertise in the field of language usage. The national reviewers also have extensive expertise in the fields of language usage standards, curriculum, and/or assessment design. The national psychometrician has experience in alignment, equating, and test analysis. Participants analyzed the alignment between the state's language usage content standards and the test questions on the assessment. The language usage content standards were used to describe the expectations for what students are to know and do. The reviewers determined the alignment of test questions to the ISAT content standards. A list of reviewers is provided below. In addition, a brief summary of each national expert's professional qualifications is provided in Appendix A. The final results of this study indicate that there is strong alignment between the Idaho Language Usage Grade 3 through 8 and 10 content standards, goals, and objectives and the ISAT Language Usage assessment.

State of Idaho Reviewers
Lisa Day
Angela Godfrey
Tracey Houston
Renda Ware

National Expert Reviewers
Carla Grasty
Janet St. John
Jean Stratton
Margaret Weldon, Ed.D.

National Psychometrician
Xiang Bo Wang, Ph.D.

Introduction

This alignment study is based on the work of Dr. Norman Webb, from the Wisconsin Center for Educational Research, University of Wisconsin–Madison, who states that the alignment of the content standards for student learning with assessments for measuring students' attainment of these expectations is an essential component for an effective standards-based education system. The study models Webb's procedures, including the use of depth-of-knowledge (DOK) categories, as well as Webb's definition of alignment (Webb 2002, p. 3). The definition is as follows:

> **Alignment** is defined as the degree to which expectations and assessments are in agreement and serve in conjunction with one another to guide the system toward students learning what they are expected to know and do. As such, alignment is a quality of the relationship between expectations and assessments and not a specific attribute of either of these two system components. Alignment describes the match between expectations and assessment that can be legitimately improved by changing either student expectations or assessments. Seen as a relationship between two or more system components, alignment can be determined by using the multiple criteria described in detail in a National Institute of Science Education (NISE) research monograph, *Criteria for Alignment of Expectations and Assessments* (Webb 2002).

The *Idaho Standards Achievement Tests* (ISAT) is a statewide, mandated testing program. The tests are given in grades 3 through 8 and 10. They consist of multiple-choice tests in the core subjects of language usage, mathematics, reading, and science. The ISAT allows teachers to measure student achievement based on Idaho's content standards. All questions are reviewed by Idaho educators for content and sensitivity.

Intensive training was provided for all reviewers to understand Webb's Alignment Model and Criteria by Dr. Margaret Weldon. They were first trained to identify the depth-of-knowledge (DOK) level for the content standards, objectives, and the test questions. This training included reviewing the definitions of the four DOK levels, as defined by Webb (2006). Training also included reviewing examples of test questions aligned to DOK and the alignment process to be followed.

Alignment Process

The alignment process is outlined in the following chart.

Idaho 2008 Alignment Process

Task 1: Determining the depth of knowledge (DOK) of each objective
Reviewers individually determined DOK for each objective. They discussed their DOK ratings in order to reach a group consensus. (See Appendix C.)

Task 2: Taking the test
Reviewers took the test, recording their answers and comments about the test questions on a separate document and/or in their test booklets.

Task 3: Determining what each test question measured and the DOK for each test question
Step 1
Using the first three test questions reviewers independently determined what each question measured by assigning it to a primary objective and a secondary objective, if applicable. A group discussion took place; however, reaching consensus on determining what each question measured was not required.

Step 2
Reviewers independently determined the DOK of the first three questions. Reviewers were instructed to code only **one** DOK (Level 1, 2, 3, or 4) for each of the three questions. Reviewers also independently noted any source of challenge for the first three questions. A group discussion took place; however, reaching consensus on the DOK of the first three questions was not required.

Step 3
Reviewers continued to independently determine the primary objective and the secondary objective, if applicable, for the remainder of the test questions.

Step 4
Reviewers independently determined the DOK for the remainder of the test questions. Again, the reviewers were instructed to code only **one** DOK for each of the remaining test questions.

Throughout the alignment process, reviewers independently noted any source of challenge for each test question, providing written comments as necessary.

Task 4: Summarizing alignment criteria of test questions
Once reviewers determined the primary and/or secondary objective for each test question and the DOK for each test question, they analyzed the entire test for:
      Depth-of-knowledge consistency
      Categorical concurrence
      Range-of-knowledge correspondence
      Balance of representation

Task 5: Debriefing Questionnaire
Reviewers independently shared feedback about the process, the test questions, and the standards and objectives.

Alignment Criteria

Reviewers independently assessed specific criteria related to the content agreement between the Idaho Language Usage content standards and the test questions on the *Idaho Standards Achievement Tests*. The four criteria receiving major attention were: DOK consistency, categorical concurrence, range-of-knowledge correspondence, and balance of representation. For each alignment criterion, an acceptable level was defined by what would be required to assure that a student had met the content standards. Along with the defined requirements, reviewers also used their professional judgment and experience in the classroom to determine whether an acceptable level for each criterion was met.

Depth-of-Knowledge Consistency

For the purpose of this study, Webb's definition of DOK consistency was used. According to Webb (2002), DOK consistency between content standards and test items indicates alignment if what is elicited from students on the test is as demanding cognitively as what students are expected to know and do as stated in the content standards. For consistency to exist between the assessment and the content standards, an item should be coded with the same DOK level as the content standards or one level above the DOK level of the content standard. Reviewers indicated "Yes" if the DOK levels of the test question and content standard were the same or one level above that of the content standards. If these were not consistent, reviewers indicated "No" and stated why. Interpreting and assigning DOK levels to content standards and test questions is an essential requirement of alignment analysis. (See Appendix D, Tables LU3.1 – LU10.1 for each grade level.)

The four levels of Depth-of-Knowledge were defined for language usage (See Appendix B.). Detailed descriptions (Webb 2006) help to clarify what the four different levels represent in language usage.

Categorical Concurrence

According to Webb (2002), an important aspect of alignment between each standard and the test is whether both address the same content categories. The categorical concurrence criterion provides a very general indication of alignment if the standards and the test incorporate the same content. For this alignment study, this criterion was judged by first allowing reviewers to make a determination as to whether the test as a whole included questions measuring content from each of the standards. The reviewers were told to use their professional opinions, as well as the Webb guiding principle to determine that at least six questions measuring content from each standard is a good indicator of categorical concurrence between the standard and the test (Webb, 2002, p. 7).

Using Webb's model, the number of questions, six, is based on estimating the number of questions that could produce a reasonably reliable subscale for estimating students' mastery of content on that subscale. Of course, many factors have to be considered in determining what a reasonable number is, including the reliability of the subscale, the mean score, and the cutoff score for determining mastery. Using a procedure developed by Subkoviak (1988), and assuming that the cutoff score is the mean and that the reliability of one item is 0.1, it was estimated that six questions would produce an agreement coefficient of at least 0.63. This indicates that about 63% of the group would be consistently classified as masters or non-masters if two equivalent test administrations were employed. The agreement coefficient would increase if the cutoff score was increased to one standard deviation from the mean to 0.77 and, with a cutoff score of 1.5 standard deviations from the mean, to 0.88.

Six questions were assumed as a minimum for an assessment measuring content knowledge related to a standard and as a basis for making some decisions about students' knowledge of that standard. If the mean for six questions is three and one standard deviation is one question, then a cutoff score set at four would produce an agreement coefficient of 0.77. Any fewer questions with a mean of one-half of the questions would require a cutoff that would only allow a student to miss one question. This would be a very stringent requirement, considering a reasonable standard error of measurement on the subscale. (See Appendix D, Tables LU3.2 – LU10.2 for each grade level.)

Range-of-Knowledge Correspondence

For standards and the test questions to be aligned, the breadth of knowledge required on both must be comparable. The range-of-knowledge criterion is used to judge whether a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer the test questions. For an acceptable range-of-knowledge, at least 50% of the objectives for a standard must have at least one related test question.
(See Appendix D, Tables LU3.3 – LU10.3 for each grade level.)

Balance of Representation

The balance of representation is met if the emphasis of content and performance supplied by the questions (primary, secondary, or both) corresponds to the standards for the test as a whole. Reviewers determined whether the test questions were distributed among the objectives of the standards that were assessed. (See Appendix D, Tables LU3.4 – LU10.4 for each grade level.)

The balance-of-representation criterion is used to indicate the degree to which one standard is given more emphasis on the assessment than another. An index is used to judge the distribution of the test questions. The index in this study is computed by considering the difference in the proportion of objectives and the proportion of hits assigned to the standard. An index value of 1 signifies perfect balance and is obtained if the hits (questions corresponding to standards) are equally distributed among the objectives. Index values that approach 0 signify that a large proportion of the hits are on only one or two of all of the objectives hit. Depending on the number of objectives and the number of hits, a unimodal distribution has an index value of less than 0.5. A bimodal distribution has an index value of around 0.55 or 0.6. Index values of 0.7 or higher indicate that questions/activities are distributed among all of the objectives at least to some degree. Index values between 0.6 and 0.7 indicate the balance-of-representation criterion has only been "weakly" met. (Table 1)

A summary of Webb's alignment criteria can be found in Table 1.

Table 1: Alignment Levels for the Four Criteria

Alignment Level	Depth-of-Knowledge Consistency	Categorical Concurrence	Range-of-Knowledge Correspondence	Balance of Representation
Yes	50%	mean is 6 or more	50%	.70
Weak	40% - 49%	mean is 5 to 5.9	40% - 49%	.60 - .69
No	less than 40%	mean is less than 5	less than 40%	less than .60

The results for each of the four criteria discussed in this section were calculated using Webb's methodology and the reviewers' averaged ratings, along with their comments. The results for depth-of-knowledge consistency, categorical concurrence, range-of-knowledge correspondence, and balance of representation are found in Appendix D, Tables LU3.1 – LU10.4 for each grade level.

Source-of-Challenge Criterion

Reviewers noted source-of-challenge issues for the test questions. The sources of challenge may include such issues as questions containing misleading factual information, questions requiring prior knowledge, questions with possible clueing among distractors, and questions deemed by the reviewer as having two possible correct answers.

Results of Alignment Analysis

There are two standards in language usage: Writing Process and Writing Components. Table 2 is the consensus of the eight reviewers' coding the DOK levels to the objectives of the standards by grade. Grades 3 and 4 have 30% of objectives coded at DOK Level 1. At least 50% of objectives are coded at DOK Level 2 for grades 3 through 8 and 10. In addition, grade 8 has the largest percentage (60%) of objectives coded at DOK Level 2. Grades 5 and 6 have at least 20% of objectives coded at DOK Level 3.

Table 2: Content Standards by Depth-of-Knowledge in Language Usage

Grade	Number of Objectives	DOK Level	Objectives by DOK Level	
			Number	*Percent*
3	13	1 2 3	4 7 2	30.77% 53.85% 15.38%
4	13	1 2 3	4 7 2	30.77% 53.85% 15.38%
5	14	1 2 3	4 8 2	28.57% 57.14% 21.43%
6	14	1 2 3	4 7 3	28.57% 50.00% 21.43%
7	14	1 2 3	4 8 2	28.57% 57.14% 14.29%
8	14	1 2 3	4 9 1	28.57% 64.29% 7.14%
10	12	1 2 3	1 9 2	8.33% 75.00% 16.67%

Reliability among Reviewers

The intra-class correlation is based on the mean squares from the analysis of variance of a two-way random effects model, reviewers crossed with items (Shrout and Fleiss 1979) as described in Appendix E. In general, an average correlation of over 0.70 is considered acceptable. However, increasing the number of reviewers may increase the reliability levels. Table 3 below provides a summary of the intra-class correlation for language usage Grades 3 through 8 and 10. In addition, the percentage of questions coded the same DOK by at least seven of the eight reviewers are also provided.

Table 3: Summary of Reliability

Grades	Number of Questions	Intra-Class Correlation	Percentage of Questions Coded the Same DOK
3	40	.91	37.5%
4	40	.86	40.0%
5	42	.90	38.1%
6	42	.91	26.2%
7	45	.90	51.1%
8	45	.89	37.8%
10	45	.81	37.8%

The intra-class-correlation range of eight reviewers is between .81 and .91. All are above .70 and are in the acceptable range. (Refer to Appendix E for the calculation modes.)

Summary of Results

The summary results of alignment analysis for Language Usage by grade and criteria are presented in Table 7. "Yes" indicates meeting the acceptable level, for depth-of-knowledge consistency, categorical concurrence, range-of-knowledge correspondence, and balance of representation.

Table 4: Summary of Alignment Results for Language Usage

Grade	Standard	Depth-of-Knowledge Consistency	Categorical Concurrence	Range-of-Knowledge Correspondence	Balance of Representation
3	3 5	Yes Yes	Yes Yes	Yes Yes	Yes Yes
4	3 5	Yes Yes	Yes Yes	Yes Yes	Yes Yes
5	3 5	Yes Yes	Yes Yes	Yes Yes	Yes Yes
6	3 5	Yes Yes	Yes Yes	Yes Yes	Yes Yes
7	3 5	Yes Yes	Yes Yes	Yes Yes	Yes Yes
8	3 5	Yes Yes	Yes Yes	Yes Yes	Yes Yes
10	3 5	Yes Yes	Yes Yes	Yes Yes	Yes Yes

Conclusion

Statistical analysis and judgments of the eight reviewers support the conclusion that there is strong alignment between the Language Usage Idaho Content Standards and objectives and the ISAT Language Usage assessments. This is indicated by the acceptable alignment levels for depth-of-knowledge consistency, categorical concurrence, range-of-knowledge correspondence, and balance of representation for both standards in grade 3 through 8 and 10.

There is adequate depth-of-knowledge consistency and categorical concurrence for both standards in grades 3 through 8 and 10. At least 50% of DOK levels coded to the questions are "at" or "above" the DOK consensus level for the objective. Categorical concurrence for both standards in grades 3 through 8 and 10 is "Yes" which indicates a mean of six or more questions reviewers coded as corresponding to a standard.

Statistical findings reveal that there is range-of-knowledge correspondence. Questions cover at least 50% of the objectives within each standard. In addition, there is a balance of representation of the language usage standards aligned to the language usage assessments. The alignment level for both standards in grades 3 through 8 and 10 is "Yes." This indicates that the test questions were equally aligned to the objectives within the standards.

In their analysis of the *Idaho Standards Achievement Tests* and the ISAT Language Usage Standards, the eight reviewers reached consensus on the DOK levels of the language usage objectives. The range of the intra-class correlation among reviewers to the test questions is 0.81 – 0.91 and that is considered acceptable. The final result of the alignment study reflects there is strong alignment between the ISAT Language Usage assessments and the language usage standards and objectives.

.

References

Shrout, P.E. and Fleiss, J.L. (1979) "Intra-class Correlations: Uses in Assessing Rater Reliability." *Psychological Bulletin*, 2, 420–428.

Subkoviak, M.J. (1988). "A Practitioner's Guide to Computation and Interpretation of Reliability Indices for Mastery Tests." *Journal of Educational Measurement*, *25*(1), 47–55.

Webb, N.L. (2002). *Alignment Study in Language Arts, Mathematics, Science, and Social Studies of State Standards and Assessments for Four States.* State Collaborative on Assessment & State Standards (SCASS).Technical Issues in Large-Scale Assessment (TILSA): University of Wisconsin, Wisconsin Center for Education Research.

Webb, N. L., Depth of Knowledge Levels. personal email, 20 November 2006.

Appendix A

Biographies of the National Experts

Carla J. Grasty

Ms. Carla J. Grasty has worked as a school improvement consultant for the Heartland Area Education Agency (AEA11). Ms. Grasty has trained teachers in AEA11 schools in the 6 + 1 Traits of Writing program, which instructs students on how to assess writing using the rubric developed for the program. She participates with other representatives from AEA11 to plan, implement, and evaluate this service area as required by Iowa code. Ms. Grasty also trains teachers in Project CRISS, a model of student-owned learning using research-based strategies to develop comprehension in content areas. She has established building projects as a Phase III building coordinator that align with district goals and curriculum improvement plans, keeping detailed records of each project. Ms. Grasty also worked with Dynamic Indicators of Basic Early Learning Skills (DIBELS), which is a research-based early intervention for basic early literacy skills, and is trained in the standardized testing techniques for DIBELS. She has participated as a reading national expert for other state alignment studies based on Dr. Norman Webb's methodology. Ms. Grasty is currently an adjunct faculty member of the psychology and child development departments at Des Moines Area Community College. In the past she was an adjunct faculty member in the psychology department at Upper Iowa University and taught G.E.D. courses at Des Moines Area Community College for four years. Also, she was an elementary school teacher for 11 years in Stuart-Menlo Elementary School and was the director and head teacher at See-Saw Preschool for 11 years. Ms. Grasty is an active member of her community; she was appointed by the governor's office to work with elected officials in the implementation of the Work Force Investment Act (2000–2005). The committee's task is to establish a new service model.

Ms. Grasty received a B.S. degree from Iowa State University, Ames, Iowa, in elementary education with an early childhood endorsement. She earned a M.S. degree from Iowa State University in elementary guidance counselor with a psychology emphasis.

Janet St. John

Ms. Janet St. John has been a freelance writer, editor, editorial consultant, book reviewer, and proofreader for 16 years, working primarily with educational publishers and literary publications. Her client list currently includes several test publishing companies including Educational Testing Service (ETS) and the American Library Association (*Booklist*). Also, Ms. St. John has participated as a reading national expert for state alignment studies based on Dr. Norman Webb's methodology. She has taught poetry at the Ragdale Foundation and creative writing at William Rainey Harper College. Ms. St. John was the founder and managing editor of the *Baybury Review,* an annual literary magazine, and is a published poet. Her poetry chapbook, *Moving Pictures*, won the University of Alaska-Fairbanks Midnight Sun Chapbook Contest and was published in 1997. Ms. St. John has published individual poems in a variety of literary magazines including *The Nebraska Review, Passages North, California Quarterly, The Spoon River Poetry Review, Poet Lore, Puerto del Sol,* and *Main Street Rag.*

She received her B.A. in Communications with a minor in English Literature from DePaul University and her M.F.A. in Writing from Vermont College.

Jean Stratton

Ms. Jean A. Stratton has worked with Re:Learning New Mexico, an organization committed to fostering educational changes so self-renewing school communities can develop and flourish. Ms. Stratton began as a senior consultant and then became director in 1997. She has continued working with Re:Learning from 2000 to the present as an educational consultant through GET, Inc. (Graphics, Editing and Technical Writing for Education). Ms. Stratton recently developed a standards implementation process for Re:Learning that included an overview of standards, the design of standards-based units, and follow-up professional development in instructional strategies. In 2001 she co-authored BEST (Building Essential Schools Together), a set of self-assessment tools that help schools develop a strategic plan. From 1997 to 2001 Ms. Stratton worked closely with Salt River Pima Maricopa Indian Community, providing systemic planning, curriculum development, and professional development for teachers. In addition, she held more than 30 workshops for the Navajo Nation Rural Systemic Initiative across the nation in curriculum design, instructional strategies, and systemic planning. In 1992 Ms. Stratton co-authored Shades of Change for Schools, a set of benchmarks that schools can use for self-assessment. She has worked on all revisions to date as well as being involved in Revitalization of School Communities through Partnerships for Student Success (Community Shades of Change).

In addition to her affiliation with Re:Learning, Ms. Stratton has conducted hundreds of workshops and given many presentations to schools, districts, and community members on standards-based curriculum design, strategic planning, team building, alternative assessment, portfolio design, coaching in instructional strategies, and Shades of Change. She has also been involved in state implementation of school reform efforts such as the Comprehensive School Reform Design program and Goals 2000. As part of the Curriculum Reform Project: Science, Mathematics, and Thinking across Disciplines, which was funded by a grant from the U.S. Department of Education, Ms. Stratton prepared a paper for the University of Colorado on how higher order thinking skills can be integrated into the general curriculum.

Ms. Stratton taught English classes to junior and senior high school students for 16 years. She received a B.A. degree in English and journalism from Eastern New Mexico University, Portales, New Mexico, and an M.A. degree in rhetoric and the teaching of composition from New Mexico State University, Las Cruces, New Mexico.

Margaret E. Weldon, Ed.D.

Dr. Margaret E. Weldon has worked as an assessment specialist for the Alabama Department of Education. She manages writing assessment program development and administration (grades 5, 7, and 10), develops writing prompts and conducts bias and content reviews. She led the development of the reading assessment (grades 3–8) for the Alabama Reading and Mathematics Test and the reading comprehension and language subject-area tests of the Alabama High School Graduation Exam (3rd ed.), as well as collaborating on the development of the Alabama Early Learning Assessment—K, 1, and 2 reading tests. Dr. Weldon has conducted statewide writing programs for teachers and administrators on composition, instructional strategies, holistic scoring, and reading instruction. She has participated in NAEP item reviews for reading and writing and in standard setting using Bookmark and Modified-Angoff methodologies. She has also participated as a reading national expert for state alignment studies directed by Dr. Norman Webb.

Dr. Weldon was a classroom teacher and administrator for 19 years in the Montgomery public schools; as a central office administrator, she directed the implementation of the state assessment program for a school system of 35,000 students. She was English department chairman when she taught secondary English. Also, Dr. Weldon was a Title 1 reading specialist.

She received a B.S. degree in secondary English education, a M.Ed. degree in secondary reading education, and an Ed.D. degree in educational leadership, foundations, and technology from Auburn University, Auburn, Alabama.

Xiang Bo Wang, Ph.D. (National Psychometrician)

Dr. Xiang Bo Wang has worked at four national educational testing and assessment institutions, dedicating himself to their statistical, psychometric, and operational excellence. At the Law School Admission Council where he served as a research scientist between 1993 and 2000, Dr. Wang played a principal role in building its entire statistical and psychometric system for test analyses, equating, and security investigation systems. At Educational Testing Service, where he was Psychometrician II between 2000 and 2001, not only did he streamline some time-consuming processes on the computer adaptive, *Test of English as a Foreign Language*, but was one of the highest rated instructors at ETS Global Institute, teaching educators from all over the world on special topics ranging from standard setting, test analysis, and equating. Between 2001 and 2003, Dr. Wang was the Director of Statistical Analysis at the Chauncey Group International where he oversaw the accuracy and efficiency of the statistical and psychometric operations of more than 100 licensure and certification examinations. In addition, he regularly conducted standard setting sessions and alignment studies, and he presented to boards of trustees of various clients.

In July 2003, Dr. Wang joined the College Board as a senior research scientist and psychometrician, and he has worked in areas of statistical infrastructure building, ethnic studies, diagnostics, alignment, standard setting, computer adaptive testing, test analysis and equating.

In addition, Dr. Wang is currently working with American College Testing (ACT) and serving as a member of the Award Committee of the National Council on Measurement in Education (NCME). Between 1982 and 1985, he also taught English as a Second Language as an assistant professor in Hunan University, China. Dr. Wang received a B.S. from Hunan University, and a Masters and Ph.D. from the University of Hawaii in Educational Psychology and Statistics.

Appendix B

Depth-of-Knowledge Levels

Depth of Knowledge ─ Language Usage

In language arts, four depth-of-knowledge levels were used to judge writing objectives and assessment tasks.

Level 1 requires the student to write or recite simple facts. The focus of this writing or recitation is not on complex synthesis or analysis, but on basic ideas. The students are asked to list ideas or words, as in a brainstorming activity, prior to written composition; are engaged in a simple spelling or vocabulary assessment; or are asked to write simple sentences. Students are expected to write, speak, and edit using the conventions of standard English. This includes using appropriate grammar, punctuation, capitalization, and spelling. Students demonstrate a basic understanding and appropriate use of such reference materials as a dictionary, thesaurus, or website. Some examples that represent, but do not constitute all of, Level 1 performance are:

- Use punctuation marks correctly.
- Identify Standard English grammatical structures, including the correct use of verb tenses.

Level 2 requires some mental processing. At this level, students are engaged in first-draft writing or brief extemporaneous speaking for a limited number of purposes and audiences. Students are expected to begin connecting ideas, using a simple organizational structure. For example, students may be engaged in note-taking, outlining, or simple summaries. Text may be limited to one paragraph. Some examples that represent, but do not constitute all of, Level 2 performance are:

- Construct or edit compound or complex sentences, with attention to correct use of phrases and clauses.
- Use simple organizational strategies to structure written work.
- Write summaries that contain the main idea of the reading selection and pertinent details.

Level 3 requires some higher-level mental processing. Students are engaged in developing compositions that include multiple paragraphs. These compositions may include complex sentence structure and may demonstrate some synthesis and analysis. Students show awareness of their audience and purpose through focus, organization, and the use of appropriate compositional elements. The use of appropriate compositional elements includes such things as addressing chronological order in a narrative, or including supporting facts and details in an informational report. At this stage, students are engaged in editing and revising to improve the quality of the composition. Some examples that represent, but do not constitute all of, Level 3 performance are:

- Support ideas with details and examples.
- Use voice appropriate to the purpose and audience.
- Edit writing to produce a logical progression of ideas.

Depth-of-Knowledge ─ Language Usage (continued)

Level 4 requires higher-level thinking. The standard at this level is a multi-paragraph composition that demonstrates the ability to synthesize and analyze complex ideas or themes. There is evidence of a deep awareness of purpose and audience. For example, informational papers include hypotheses and supporting evidence. Students are expected to create compositions that demonstrate a distinct voice and that stimulate the reader or listener to consider new perspectives on the addressed ideas and themes. An example that represents, but does not constitute all of, Level 4 performance is:

- Write an analysis of two selections, identifying the common theme and generating a purpose that is appropriate for both.

(Webb 2006)

Appendix C

Depth-of-Knowledge Consensus

Depth-of-Knowledge Consensus
Language Usage
Grade 3

Standard *Goal* Objective	Depth of Knowledge of the Objectives **Level 1** **Level 2** **Level 3**
	Group Consensus
3: Writing Process	
3.1: Acquire Prewriting Skills	
3.1.2 Generate the main idea.	2
3.1.3 Plan writing using organizational strategies (e.g., graphic organizer, chart).	3
3.1.4 Identify an appropriate writing format for purpose and audience.	2
3.3: Acquire Skills for Revising a Draft	
3.3.1 Revise draft for meaning.	3
3.3.2 Revise draft by adding details to enhance audience understanding.	2
3.3.3 Identify words and sentences that need to be rearranged to clarify meaning.	2
3.4: Acquire Skills for Editing a Draft	
3.4.1 Edit the draft using a simple editing checklist.	2
5: Writing Components	
5.2: Acquire Spelling Skills	
5.2.1 Spell correctly Grade 3 high-frequency words.	1
5.2.3 Apply spelling rules appropriate to grade level to spell accurately.	2
5.3: Acquire Skills for Sentence Structure	
5.3.1 Identify and use three types of sentences (exclamatory, declarative, and interrogative).	2
5.3.2 Use past and present verb tenses, including irregular verbs.	1

5.4: Acquire Skills for Using Conventions	
5.4.1 Capitalize proper nouns, titles, and holidays.	1
5.4.2 Use commas in: • series • dates • addresses • letters	1

Depth-of-Knowledge Consensus
Language Usage
Grade 4

Standard *Goal* Objective	Depth of Knowledge of the Objectives **Level 1** **Level 2** **Level 3**
	Group Consensus
3: Writing Process	
3.1: Acquire Prewriting Skills	
3.1.2 Generate the main idea.	2
3.1.3 Use organizational strategies appropriate for writing.	2
3.1.4 Select an appropriate writing format for purpose and audience.	3
3.3: Acquire Skills for Revising a Draft	
3.3.1 Revise draft for meaning and clarity	3
3.3.2 Revise draft by adding details to enhance audience understanding.	2
3.3.4 Rearrange words and sentences as needed to clarify meaning.	2
3.4: Acquire Skills for Editing a Draft	
3.4.1 Edit the draft using an editing checklist with common editing marks	2
5: Writing Components	
5.2: Acquire Spelling Skills	
5.2.1 Spell correctly Grade 4 high-frequency words and common content area (e.g., science, social studies) words.	1
5.2.3 Apply spelling rules appropriate to grade level to spell accurately.	2
5.3: Acquire Skills for Sentence Structure	
5.3.1 Use simple and complex sentences.	2
5.3.2 Identify: • future verb tenses • adjectives • personal pronouns • conjunctions	1

5.4: Acquire Skills for Using Conventions	
5.4.1 Correctly punctuate and capitalize titles, books, geographical names.	1
5.4.2 Identify comma use in a direct address ("John, come here.") and in compound sentences.	1

Depth-of-Knowledge Consensus
Language Usage
Grade 5

Standard *Goal* Objective	Depth of Knowledge of the Objectives Level 1 Level 2 Level 3
	Group Consensus
3: Writing Process	
3.1: Acquire Prewriting Skills	
3.1.2 Generate a main idea appropriate to the type of writing.	2
3.1.3 Select organizational strategies appropriate for writing.	2
3.1.4 Select an appropriate writing format for purpose and audience.	3
3.3: Acquire Skills for Revising a Draft	
3.3.1 Revise draft for meaning, clarity, and effective sequencing.	3
3.3.2 Revise draft by adding details to enhance audience understanding.	2
3.3.3 Apply and add transition words to clarify sequence.	2
3.3.4 Rearrange words, sentences, and paragraphs as needed to clarify meaning.	2
3.4: Acquire Skills for Editing a Draft	
3.4.1 Edit the draft using an editing checklist with common editing marks.	2
5: Writing Components	
5.2: Acquire Spelling Skills	
5.2.1 Spell correctly Grade 5 high-frequency words and content area words.	1
5.2.3 Apply spelling rules appropriate to grade level to spell accurately.	2
5.3: Acquire Skills for Sentence Structure	
5.3.1 Identify complex sentences with subject and verb agreement.	2
5.3.2 Use correctly: • future verb tenses • adjectives • personal pronouns • conjunctions • adverbs	1

5.4: Acquire Skills for Using Conventions	
5.4.1 Apply capitalization correctly in writing.	1
5.4.2 Identify a colon to introduce a list and in a greeting. Identify quotation marks to punctuate dialogue.	1

Depth-of-Knowledge Consensus
Language Usage
Grade 6

Standard *Goal* Objective	Depth of Knowledge of the Objectives Level 1 Level 2 Level 3
	Group Consensus
3: Writing Process	
3.1: Acquire Prewriting Skills	
3.1.2 Generate a main idea appropriate to the type of writing.	2
3.1.3 Select organizational strategies appropriate for writing.	2
3.1.4 Apply an appropriate writing format for purpose and audience.	3
3.3: Acquire Skills for Revising a Draft	
3.3.1 Revise draft for meaning, clarity and effective sequencing.	3
3.3.2 Add details to more effectively accomplish the purpose of writing.	3
3.3.3 Apply and add transition words to clarify sequence.	2
3.3.4 Rearrange words, sentences, and paragraphs to enhance writing style.	2
3.4: Acquire Skills for Editing a Draft	
3.4.1 Edit the draft using an editing checklist with common editing marks.	2
5: Writing Components	
5.2: Acquire Spelling Skills	
5.2.1 Spell correctly Grade 6 high-frequency words and content area words.	1
5.2.3 Apply spelling rules appropriate to grade level (e.g., less common prefixes, suffixes and plurals) to spell accurately.	2
5.3: Acquire Skills for Sentence Structure	
5.3.1 Identify complex sentences with subject and verb agreement.	2
5.3.2 Use correctly: • future verb tenses • adjectives • personal pronouns • conjunctions • adverbs	1

5.4: Acquire Skills for Using Conventions	
5.4.1 Apply capitalization correctly in writing.	1
5.4.2 Use quotation marks and commas to punctuate dialogue.	1

Depth-of-Knowledge Consensus
Language Usage
Grade 7

Standard *Goal* Objective	Depth of Knowledge of the Objectives Level 1 Level 2 Level 3
	Group Consensus
3: Writing Process	
3.1: Acquire Prewriting Skills	
3.1.2 Generate a main idea or thesis appropriate to the writing.	2
3.1.3 Apply appropriate organizational strategies to plan writing.	2
3.1.4 Match appropriate writing format to purpose and audience.	3
3.3: Acquire Skills for Revising a Draft	
3.3.1 Revise draft for meaning, clarity, and effective organization.	3
3.3.2 Add details and delete irrelevant or redundant information.	2
3.3.3 Arrange transition words and phrases in draft to clarify meaning and improve organization.	2
3.3.4 Apply a variety of sentence structures to improve sentence fluency and enhance writing style.	2
3.4: Acquire Skills for Editing a Draft	
3.4.1 Edit the draft using an editing checklist with common editing marks.	2
5: Writing Components	
5.2: Acquire Spelling Skills	
5.2.1 Spell correctly Grade 7 high-frequency words and content area words.	1
5.2.3 Apply spelling rules appropriate to grade level (e.g., less common prefixes, suffixes and plurals) to spell accurately.	2
5.3: Acquire Skills for Sentence Structure	
5.3.1 Use four types of sentences correctly (exclamatory, declarative, interrogative, and imperative).	2
5.3.2 Use correctly: • future verb tenses • adjectives • personal pronouns • conjunctions • adverbs	1

5.4: Acquire Skills for Using Conventions	
5.4.1 Apply capitalization correctly in writing.	1
5.4.2 Use commas, including in appositives; use parentheses and semicolon in writing.	1

Depth-of-Knowledge Consensus
Language Usage
Grade 8

Standard *Goal* Objective	Depth of Knowledge of the Objectives **Level 1** **Level 2** **Level 3**
	Group Consensus
3: Writing Process	
3.1: Acquire Prewriting Skills	
3.1.2 Generate a main idea or thesis appropriate to the writing.	2
3.1.3 Apply appropriate organizational strategies to plan writing.	2
3.1.4 Match appropriate writing format to purpose and audience.	2
3.3: Acquire Skills for Revising a Draft	
3.3.1 Revise draft for meaning, clarity, and effective organization.	3
3.3.2 Add details and delete irrelevant or redundant information.	2
3.3.3 Arrange transition words and phrases in draft to clarify meaning and improve organization.	2
3.3.4 Apply a variety of sentence structures to improve sentence fluency and enhance writing style.	2
3.4: Acquire Skills for Editing a Draft	
3.4.1 Edit the draft using an editing checklist with common editing marks.	2
5: Writing Components	
5.2: Acquire Spelling Skills	
5.2.1 Spell correctly Grade 8 high-frequency words and content area words.	1
5.2.3 Apply spelling rules appropriate to grade level (e.g., less common prefixes, suffixes and plurals) to spell accurately.	2
5.3: Acquire Skills for Sentence Structure	
5.3.1 Write correct and varied sentence structure (e.g., simple, compound, complex, and compound-complex).	2
5.3.2 Use correctly: • future verb tenses • adjectives • personal pronouns • conjunctions • adverbs	1

5.4: Acquire Skills for Using Conventions	
5.4.1 Apply capitalization correctly in writing.	1
5.4.2 Use commas, including in appositives; use parentheses and semicolon.	1

Depth-of-Knowledge Consensus
Language Usage
Grade 10

Standard *Goal* Objective	Depth of Knowledge of the Objectives **Level 1** **Level 2** **Level 3**
	Group Consensus
3: Writing Process	
3.1: Acquire Prewriting Skills	
3.1.2 Generate a main idea or thesis appropriate to a type of writing.	2
3.1.3 Apply organizational strategies to plan writing.	2
3.1.4 Match format to purpose and audience.	2
3.3: Acquire Skills for Revising a Draft	
3.3.1 Revise draft for meaning, clarity, and effective organization.	3
3.3.2 Add relevant details and delete irrelevant or redundant information.	2
3.3.3 Use transitional words and phrases to clarify meaning and improve organization.	2
3.3.4 Use a variety of sentence structures to improve sentence fluency and enhance style.	2
3.4: Acquire Skills for Editing a Draft	
3.4.2 Edit for correct punctuation, spelling, grammar, and usage errors.	2
5: Writing Components	
5.3: Acquire Skills for Sentence Structure	
5.3.1 Apply correct and varied sentence types in writing.	2
5.3.2 Edit for agreement, word usage, parallel structure, and fluency.	3

5.4: Acquire Skills for Using Conventions	
5.4.1 Demonstrate in writing the correct use of conventions emphasizing pronoun/antecedent agreement, subject/verb agreement, adjective/adverb usage, verb tense, verbals, appositives, compound-complex sentences, clauses, and parallel structure	2
5.4.2 Demonstrate in writing the correct use of punctuation and capitalization.	1

Appendix D

Results of the Alignment Analysis

Grade 3

Table LU3.1: Summary of Depth-of-Knowledge Consistency

Standard	Goals	Hits		Percent of Questions at DOK Level						DOK Consistency
		Mean	SD	Under		At		Above		
				Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	7	27	4.54	22.11	6.74	61.38	7.17	16.51	8.39	Yes
5. Writing Components	6	20.5	2.93	10.80	5.84	73.62	6.33	15.58	9.70	Yes
Total	13	47.5	4.60	16.45	8	67.5	8.65	16.04	.66	

Table LU3.2: Summary of Categorical Concurrence

Standard	Goals	Level of Objectives			Hits		Categorical Concurrence
		Level	Number of Objectives	Percentage of Objectives	Mean	SD	
3. Writing Process	7	1	0	0%	27	4.54	Yes
		2	5	71.43%			
		3	2	28.57%			
5. Writing Components	6	1	4	66.67%	20.5	2.93	Yes
		2	2	33.33%			
		3	0	0%			
Total	13	1	4	30.77%	47.5	4.60	
		2	7	53.85%			
		3	2	15.38%			

Grade 3 (continued)

Table LU3.3: Summary of Range-of-Knowledge Correspondence

Standard	Goals	Hits		Range of Objectives				Range of Knowledge
				# of Objectives Hit		% of Total		
		Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	7	27	4.54	6.88	.35	98.21%	.05	Yes
5. Writing Components	6	20.5	2.93	5.63	.52	93.75%	.09	Yes
Total	13	47.5	4.60	6.25	.88	95.98%	.03	

Table LU3.4: Summary of Balance of Representation

Standard	Goals	Percentage of Total Hits		Index		Balance of Representation
		Mean	SD	Mean	SD	
3. Writing Process	7	56.67%	.05	.76	.07	Yes
5. Writing Components	6	43.33%	.05	.75	.08	Yes
Total	13	50%	.09	.76	.01	

Grade 4

Table LU 4.1: Summary of Depth-of-Knowledge Consistency

Standard	Goals	Hits		Percent of Questions at DOK Level						DOK Consistency
				Under		At		Above		
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	7	24.63	6.44	20.54	7.77	70.36	11.28	9.10	7.06	Yes
5. Writing Components	6	23.5	2.20	9.79	10.65	73.64	12.94	16.56	14.17	Yes
Total	13	48.13	0.80	15.17	7.60	72.00	2.32	12.83	5.28	

Table LU4.2: Summary of Categorical Concurrence

Standard	Goals	Level of Objectives			Hits		Categorical Concurrence
		Level	Number of Objectives	Percentage of Objectives	Mean	SD	
3. Writing Process	7	1 2 3	0 5 2	0% 71.43% 28.57%	24.63	6.44	Yes
5. Writing Components	6	1 2 3	4 2 0	66.67% 33.33% 0%	23.5	2.20	Yes
Total	13	1 2 3	4 7 2	30.77% 53.85% 15.38%	48.13	0.80	

Grade 4 (continued)

Table LU4.3: Summary of Range-of-Knowledge Correspondence

Standard	Goals	Hits		Range of Objectives				Range of Knowledge
				# of Objectives Hit		% of Total		
		Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	7	24.63	6.44	6.25	0.71	89.29%	0.10	Yes
5. Writing Components	6	23.5	2.20	5.75	0.46	95.83%	0.08	Yes
Total	13	48.13	0.80	6	0.35	92.56%	0.05	

Table LU4.4: Summary of Balance of Representation

Standard	Goals	Percentage of Total Hits		Index		Balance of Representation
		Mean	SD	Mean	SD	
3. Writing Process	7	50.52%	0.08	0.77	0.08	Yes
5. Writing Components	6	49.48%	0.08	0.76	0.07	Yes
Total	13	50.00%	0.01	0.77	0.01	

Grade 5

Table LU5.1: Summary of Depth-of-Knowledge Consistency

Standard	Goals	Hits		Percent of Questions at DOK Level						DOK Consistency
		Mean	SD	Under		At		Above		
				Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	27.5	3.55	18.44	10.33	70.83	20.53	10.73	12.23	Yes
5. Writing Components	6	20.62	2.77	6.53	8.03	73.57	14.93	19.90	11.29	Yes
Total	14	48.12	4.86	12.48	8.42	72.20	1.93	15.32	6.49	

Table LU5.2: Summary of Categorical Concurrence

Standard	Goals	Level of Objectives			Hits		Categorical Concurrence
		Level	Number of Objectives	Percentage of Objectives	Mean	SD	
3. Writing Process	8	1	0	0%	27.5	3.55	Yes
		2	6	75%			
		3	2	25%			
5. Writing Components	6	1	4	66.67%	20.62	2.77	Yes
		2	2	33.33%			
		3	0	0%			
Total	14	1	4	28.57%	48.12	4.86	
		2	8	57.14%			
		3	2	21.43%			

Grade 5 (continued)

Table LU5.3: Summary of Range-of-Knowledge Correspondence

Standard	Goals	Hits		Range of Objectives				Range of Knowledge
				# of Objectives Hit		% of Total		
		Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	27.5	3.55	7.875	0.35	98.44%	0.04	Yes
5. Writing Components	6	20.62	2.77	5.375	0.52	89.58%	0.09	Yes
Total	14	48.12	4.86	6.625	1.77	94.01%	0.06	

Table LU5.4: Summary of Balance of Representation

Standard	Goals	Percentage of Total Hits		Index		Balance of Representation
		Mean	SD	Mean	SD	
3. Writing Process	8	57.10%	0.05	0.80	0.07	Yes
5. Writing Components	6	42.90%	0.05	0.84	0.06	Yes
Total	14	50.00%	0.10	0.82	0.03	

Grade 6

Table LU6.1: Summary of Depth-of-Knowledge Consistency

| Standard | Goals | Hits | | Percent of Questions at DOK Level | | | | | | DOK Consistency |
| | | Mean | SD | Under | | At | | Above | | |
				Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	28.88	3.80	27.27	12.30	62.34	16.51	10.39	6.76	Yes
5. Writing Components	6	22.88	3	5.46	4.23	59.02	16.32	35.52	15.71	Yes
Total	14	51.75	4.24	16.37	0.15	60.68	0.02	22.95	0.18	

Table LU6.2: Summary of Categorical Concurrence

| Standard | Goals | Level of Objectives | | | Hits | | Categorical Concurrence |
		Level	Number of Objectives	Percentage of Objectives	Mean	SD	
3. Writing Process	8	1	0	0%	28.88	3.80	Yes
		2	5	62.5%			
		3	3	37.5%			
5. Writing Components	6	1	4	66.67%	22.88	3	Yes
		2	2	33.33%			
		3	0	0%			
Total	14	1	4	28.57%	51.75	4.24	
		2	7	50%			
		3	3	21.43%			

Grade 6 (continued)

Table LU6.3: Summary of Range-of-Knowledge Correspondence

Standard	Goals	Hits		Range of Objectives				Range of Knowledge
				# of Objectives Hit		% of Total		
		Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	28.88	3.80	7.88	0.35	98.44%	0.04	Yes
5. Writing Components	6	22.88	3	5.75	0.46	95.83%	0.08	Yes
Total	14	51.75	4.24	6.81	1.50	97.14%	0.02	

Table LU6.4: Summary of Balance of Representation

Standard	Goals	Percentage of Total Hits		Index		Balance of Representation
		Mean	SD	Mean	SD	
3. Writing Process	8	55.78%	0.03	0.81	0.04	Yes
5. Writing Components	6	44.22%	0.03	0.77	0.04	Yes
Total	14	50.00%	0.08	0.79	0.03	

Grade 7

Table LU7.1: Summary of Depth-of-Knowledge Consistency

Standard	Goals	Hits		Percent of Questions at DOK Level						DOK Consistency
				Under		At		Above		
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	29.63	3.16	18.09	6.83	69.79	10.46	12.12	6.30	Yes
5. Writing Components	6	22.38	2	6.48	6.08	72.45	8.64	21.07	9.78	Yes
Total	14	52	5.13	12.29	8.21	71.12	1.88	16.59	6.33	

Table LU7.2: Summary of Categorical Concurrence

Standard	Goals	Level of Objectives			Hits		Categorical Concurrence
		Level	Number of Objectives	Percentage of Objectives	Mean	SD	
3. Writing Process	8	1	0	0%	29.63	3.16	Yes
		2	6	75%			
		3	2	25%			
5. Writing Components	6	1	4	66.67%	22.38	2	Yes
		2	2	33.33%			
		3	0	0%			
Total	14	1	4	28.57%	52	5.13	
		2	8	57.14%			
		3	2	14.29%			

Grade 7 (continued)

Table LU7.3: Summary of Range-of-Knowledge Correspondence

Standard	Goals	Hits		Range of Objectives				Range of Knowledge
				# of Objectives Hit		% of Total		
		Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	29.63	3.16	7.88	0.35	98.44%	0.04	Yes
5. Writing Components	6	22.38	2	5.63	0.52	93.75%	0.09	Yes
Total	14	52	5.13	6.75	1.59	96.09%	0.03	

Table LU7.4: Summary of Balance of Representation

Standard	Goals	Percentage of Total Hits		Index		Balance of Representation
		Mean	SD	Mean	SD	
3. Writing Process	8	56.91%	0.04	0.82	0.07	Yes
5. Writing Components	6	43.09%	0.04	0.78	0.07	Yes
Total	14	50%	0.10	0.80	0.02	

Grade 8

Table LU8.1: Summary of Depth-of-Knowledge Consistency

Standard	Goals	Hits		Percent of Questions at DOK Level						DOK Consistency
				Under		At		Above		
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	34.25	4.30	9.45	9.32	74.31	16.78	16.24	13.28	Yes
5. Writing Components	6	21.63	3.25	6.92	5.68	64.47	12.97	28.60	13.21	Yes
Total	14	55.88	8.93	8.19	1.79	69.39	6.96	22.42	8.74	

Table LU8.2: Summary of Categorical Concurrence

Standard	Goals	Level of Objectives			Hits		Categorical Concurrence
		Level	Number of Objectives	Percentage of Objectives	Mean	SD	
3. Writing Process	8	1	0	0%	34.25	4.30	Yes
		2	7	87.5%			
		3	1	12.5%			
5. Writing Components	6	1	4	66.67%	21.63	3.25	Yes
		2	2	33.33%			
		3	0	0%			
Total	14	1	4	28.57%	55.88	8.93	
		2	9	64.29%			
		3	1	7.14%			

Grade 8 (continued)

Table LU8.3: Summary of Range-of-Knowledge Correspondence

Standard	Goals	Hits		Range of Objectives				Range of Knowledge
				# of Objectives Hit		% of Total		
		Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	34.25	4.30	7.75	0.46	96.88%	0.06	Yes
5. Writing Components	6	21.63	3.25	6	0	100%	0	Yes
Total	14	55.88	8.93	6.88	1.24	98.44%	0.02	

Table LU8.4: Summary of Balance of Representation

Standard	Goals	Percentage of Total Hits		Index		Balance of Representation
		Mean	SD	Mean	SD	
3. Writing Process	8	61.23%	0.06	0.85	0.04	Yes
5. Writing Components	6	38.77%	0.06	0.82	0.02	Yes
Total	14	50%	0.16	0.84	0.03	

Grade 10

Table LU10.1: Summary of Depth-of-Knowledge Consistency

Standard	Goals	Hits		Percent of Questions at DOK Level						DOK Consistency
		Mean	SD	Under		At		Above		
				Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	37.13	1.89	13.36	7.75	71.26	14.62	15.38	12.38	Yes
5. Writing Components	4	20	3.70	27.83	6.74	56.35	12.38	15.82	8.36	Yes
Total	12	57.13	12.11	20.59	10.23	63.81	10.55	15.60	0.31	

Table LU10.2: Summary of Categorical Concurrence

Standard	Goals	Level of Objectives			Hits		Categorical Concurrence
		Level	Number of Objectives	Percentage of Objectives	Mean	SD	
3. Writing Process	8	1	0	0%	37.13	1.89	Yes
		2	7	87.5%			
		3	1	12.5%			
5. Writing Components	4	1	1	25%	20	3.70	Yes
		2	2	50%			
		3	1	25%			
Total	12	1	1	8.33%	57.13	12.11	
		2	9	75%			
		3	2	16.67%			

Grade 10 (continued)

Table LU10.3: Summary of Range-of-Knowledge Correspondence

Standard	Goals	Hits		Range of Objectives				Range of Knowledge
				# of Objectives Hit		% of Total		
		Mean	SD	Mean	SD	Mean	SD	
3. Writing Process	8	37.13	1.89	8	0	100%	0	Yes
5. Writing Components	4	20	3.70	4	0	100%	0	Yes
Total	12	57.13	12.11	6.00	2.83	100%	0	

Table LU10.4: Summary of Balance of Representation

Standard	Goals	Percentage of Total Hits		Index		Balance of Representation
		Mean	SD	Mean	SD	
3. Writing Process	8	65.23%	0.05	0.82	0.07	Yes
5. Writing Components	4	34.77%	0.05	0.89	0.06	Yes
Total	12	50%	0.22	0.85	0.04	

Appendix E

DOK Levels by Item and Reviewers and

Results of Intra-Class Correlation

Table LU3.5
DOK Levels by Item and Reviewers
ISAT Language Usage Grade 3

Item No.	R1	R2	R3	R4	R5	R6	R7	R8
1	3	2	2	2	2	2	2	2
2	2	2	2	1	1	2	1	1
3	3	3	2	3	3	3	3	3
4	2	2	2	2	2	3	2	2
5	2	2	2	2	2	2	2	3
6	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	2
8	1	1	1	2	1	1	1	1
9	1	1	1	1	1	1	1	1
10	2	1	1	1	2	1	1	2
11	1	2	2	2	2	1	1	1
12	3	2	2	2	2	2	2	2
13	2	2	1	2	2	2	2	2
14	2	2	2	2	2	2	2	3
15	2	1	2	1	1	1	2	2
16	3	2	3	3	2	3	3	3
17	2	2	1	2	2	2	2	3
18	2	3	2	2	3	2	3	3
19	2	2	2	2	3	2	2	3
20	2	2	3	2	3	3	2	2
21	2	2	2	1	1	2	1	2
22	1	1	1	1	1	1	1	2
23	3	2	1	2	2	2	2	2
24	2	1	2	2	1	2	2	2
25	2	1	1	1	1	1	2	2
26	3	2	2	2	2	2	2	2
27	2	1	1	1	1	1	2	2
28	2	2	1	3	1	2	2	2
29	3	2	3	3	2	3	2	3
30	3	2	2	3	2	3	2	2
31	2	2	3	3	2	2	2	2
32	3	3	3	3	3	2	3	3
33	2	2	1	3	2	3	2	2
34	1	1	1	2	1	1	1	1
35	2	1	1	2	2	2	2	1
36	2	2	1	1	2	1	1	2
37	2	1	2	1	1	2	2	2
38	1	2	2	2	1	1	2	2
39	1	1	1	2	1	1	1	2
40	2	2	2	2	1	2	2	1

Table LU4.5
DOK Levels by Item and Reviewers
ISAT Language Usage Grade 4

Item No.	R1	R2	R3	R4	R5	R6	R7	R8
1	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2
3	2	2	2	2	2	2	2	2
4	2	2	2	2	2	2	2	3
5	2	2	2	2	1	2	2	2
6	2	2	1	2	1	1	2	2
7	1	1	2	1	1	2	3	2
8	2	2	2	2	2	3	3	2
9	1	1	1	2	1	1	2	2
10	1	1	1	1	1	1	1	1
11	2	1	1	2	1	1	1	1
12	2	1	2	1	1	1	1	1
13	1	1	1	1	1	1	2	1
14	1	3	2	1	3	2	1	2
15	2	2	2	1	2	2	2	2
16	1	1	1	1	2	2	1	2
17	2	2	3	2	3	2	2	2
18	2	2	2	2	2	2	2	2
19	2	2	2	2	2	2	2	2
20	2	2	2	3	3	2	2	2
21	3	2	3	2	2	2	2	2
22	3	2	3	2	2	3	2	2
23	2	2	2	2	1	2	2	2
24	2	2	2	1	1	2	2	2
25	2	2	3	2	2	2	2	2
26	2	2	2	2	1	2	2	2
27	1	1	1	1	1	1	1	1
28	1	1	2	1	1	1	1	2
29	2	1	2	2	1	2	1	2
30	3	1	2	2	2	2	1	2
31	2	1	2	1	1	2	1	1
32	3	3	3	3	2	2	2	2
33	2	2	2	2	2	2	2	2
34	2	2	2	2	2	3	2	3
35	3	2	3	2	2	3	2	2
36	2	2	2	2	1	3	1	2
37	1	1	1	1	1	2	1	2
38	1	1	1	1	1	1	1	2
39	2	1	2	1	1	1	1	2
40	1	2	2	1	2	2	2	2

Table LU5.5
DOK Levels by Item and Reviewers
ISAT Language Usage Grade 5

Item No.	R1	R2	R3	R4	R5	R6	R7	R8
1	2	1	2	1	1	2	2	2
2	2	2	1	1	1	1	2	2
3	2	2	2	2	1	2	3	2
4	2	3	2	2	3	2	2	3
5	2	3	2	2	3	2	2	3
6	2	2	2	2	2	2	2	2
7	2	2	2	2	1	2	2	2
8	2	2	2	2	2	3	3	3
9	2	2	2	2	3	2	2	2
10	3	2	3	3	3	2	2	2
11	3	2	3	3	3	3	3	3
12	1	1	1	1	1	2	1	2
13	2	2	2	1	2	2	2	2
14	1	2	1	2	1	1	2	1
15	1	1	1	1	1	2	1	1
16	2	2	1	2	1	2	1	2
17	2	3	2	2	1	2	3	2
18	2	2	2	3	2	2	2	2
19	2	2	2	3	3	2	2	2
20	2	2	2	2	1	2	2	2
21	2	2	2	3	3	2	2	2
22	3	3	2	3	3	3	2	3
23	2	2	3	2	3	2	2	2
24	1	2	2	2	2	1	2	2
25	1	2	1	1	1	1	1	1
26	2	1	1	1	1	1	1	1
27	2	2	2	2	2	1	1	2
28	1	1	1	1	1	1	1	2
29	2	2	2	1	1	1	2	2
30	1	2	1	1	1	1	2	2
31	2	3	1	2	2	2	1	2
32	2	3	2	2	1	2	2	2
33	2	2	2	2	2	2	2	3
34	2	2	2	2	2	2	2	2
35	2	2	2	2	3	2	2	3
36	2	2	2	2	3	3	2	2
37	1	1	1	2	1	1	1	1
38	2	1	1	1	2	2	2	2
39	1	1	1	1	1	1	1	1
40	1	1	1	1	1	2	1	1
41	1	1	1	1	1	1	2	2
42	2	2	1	1	1	1	2	2

Table LU6.5
DOK Levels by Item and Reviewers
ISAT Language Usage Grade 6

Item No.	R1	R2	R3	R4	R5	R6	R7	R8
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	2
3	1	1	1	2	1	1	1	2
4	2	2	2	3	3	2	2	3
5	2	2	2	2	2	2	2	2
6	2	2	2	3	2	3	2	2
7	3	2	2	3	3	2	3	2
8	2	2	2	3	2	2	3	3
9	2	2	3	2	3	2	3	3
10	2	3	2	3	2	3	3	3
11	2	2	3	3	3	2	2	2
12	2	1	1	2	1	1	1	2
13	1	1	2	2	1	1	1	2
14	1	2	1	2	2	2	2	2
15	1	2	2	2	1	2	2	2
16	2	2	1	2	3	1	2	2
17	1	3	2	2	2	1	1	2
18	1	1	1	2	2	1	1	2
19	3	3	2	2	2	3	2	3
20	2	2	2	2	2	2	2	2
21	3	3	2	2	3	2	2	3
22	3	2	2	2	3	2	2	3
23	2	2	3	2	2	3	2	2
24	2	2	2	3	2	2	2	2
25	1	1	1	1	1	1	1	1
26	2	1	1	2	2	1	1	2
27	2	3	2	2	3	2	2	2
28	2	2	2	2	2	2	2	2
29	3	2	2	3	2	3	2	2
30	3	3	2	3	2	3	3	2
31	3	2	2	2	2	2	2	3
32	3	2	3	2	3	2	2	3
33	2	2	2	2	2	2	2	2
34	3	3	3	3	3	3	3	2
35	2	2	2	2	1	2	2	2
36	2	2	3	3	3	3	3	2
37	1	1	1	2	1	1	1	2
38	1	2	1	1	1	1	2	2
39	1	1	1	1	1	1	1	1
40	2	2	2	3	2	2	1	1
41	2	2	2	2	2	1	2	1
42	2	1	1	2	1	1	2	1

Table LU7.5
DOK Levels by Item and Reviewers
ISAT Language Usage Grade 7

Item No.	R1	R2	R3	R4	R5	R6	R7	R8
1	1	1	1	1	1	1	1	1
2	2	2	2	1	1	2	2	2
3	2	2	2	2	2	3	2	2
4	2	2	2	2	2	3	2	3
5	2	2	2	2	2	2	2	2
6	2	2	2	2	2	2	2	2
7	3	2	2	2	1	2	3	2
8	2	3	2	2	2	2	3	3
9	2	2	2	3	2	2	2	2
10	2	2	3	2	3	2	2	2
11	2	2	2	3	3	2	2	3
12	2	1	1	2	1	2	2	2
13	1	1	1	2	1	1	1	1
14	1	2	2	2	1	1	1	1
15	2	2	2	2	2	2	2	2
16	1	1	1	1	1	1	1	1
17	1	2	1	2	1	1	2	2
18	2	2	1	2	2	2	2	2
19	2	2	1	2	2	2	2	2
20	2	2	3	2	3	2	2	2
21	3	2	2	3	2	3	2	3
22	2	2	3	2	2	2	2	2
23	2	2	2	2	1	2	2	2
24	2	2	2	2	2	2	3	2
25	3	3	2	3	2	2	2	2
26	2	2	3	2	2	2	2	2
27	1	1	1	1	1	1	1	1
28	1	1	1	3	1	2	1	2
29	2	1	1	2	2	2	1	2
30	2	1	1	1	1	1	1	1
31	1	1	1	1	1	1	1	2
32	2	3	1	2	2	3	2	2
33	2	3	2	2	3	2	2	2
34	2	2	2	2	2	2	2	2
35	2	2	2	2	3	2	2	2
36	2	3	2	2	3	2	3	2
37	2	2	2	2	2	2	2	2
38	2	2	3	3	3	2	3	2
39	3	3	3	3	3	2	2	3
40	2	2	2	2	2	2	2	2
41	2	2	1	2	2	1	1	2
42	1	1	1	1	1	1	1	1
43	2	2	1	1	3	1	1	2
44	1	1	1	1	1	1	1	2
45	2	1	1	2	1	2	2	2

Table LU8.5
DOK Levels by Item and Reviewers
ISAT Language Usage Grade 8

Item No.	R1	R2	R3	R4	R5	R6	R7	R8
1	1	1	1	1	1	1	1	2
2	1	1	1	2	2	2	2	2
3	2	2	2	2	3	2	2	2
4	1	2	3	2	3	2	2	2
5	2	3	3	3	2	3	2	2
6	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	2	2
8	2	2	2	2	2	2	2	2
9	2	2	2	3	3	3	2	3
10	2	2	2	2	2	2	2	2
11	3	2	2	2	3	3	2	2
12	2	2	1	2	1	2	2	2
13	2	3	2	3	3	3	3	2
14	2	2	3	2	3	3	2	2
15	2	2	1	2	2	3	2	2
16	1	1	1	1	1	1	1	1
17	2	2	1	2	2	1	1	2
18	2	2	2	3	2	1	2	2
19	3	2	2	2	3	3	2	2
20	2	2	2	2	2	3	2	2
21	2	2	2	2	3	2	3	2
22	2	2	2	2	1	2	2	2
23	2	2	2	2	1	2	2	2
24	3	2	2	3	3	3	2	2
25	2	2	1	2	1	2	2	2
26	3	2	3	3	3	3	2	3
27	3	3	2	3	3	3	3	3
28	2	3	3	3	3	3	2	2
29	2	1	1	2	1	1	2	2
30	1	1	1	1	1	1	1	2
31	2	2	1	2	1	1	2	2
32	1	2	1	1	1	1	1	1
33	1	1	1	2	1	1	1	2
34	2	2	2	2	2	2	2	2
35	2	2	2	2	2	3	2	2
36	2	2	2	2	2	2	2	2
37	2	2	2	3	3	3	2	2
38	3	2	2	2	2	2	2	2
39	3	2	2	2	3	3	2	2
40	3	3	2	3	2	2	2	2
41	1	1	2	2	2	1	1	2
42	1	1	1	2	2	1	2	2
43	1	2	1	2	2	1	2	2
44	1	1	1	2	2	2	1	2
45	1	1	1	2	1	1	1	2

Table LU10.5
DOK Levels by Item and Reviewers
ISAT Language Usage Grade 10

Item No.	R1	R2	R3	R4	R5	R6	R7	R8
1	2	2	2	2	2	2	2	2
2	2	2	3	2	2	2	2	2
3	2	2	2	3	2	2	3	3
4	2	2	2	2	3	2	2	2
5	2	2	2	2	2	2	2	2
6	2	2	2	2	3	2	2	2
7	2	2	1	2	2	2	2	2
8	3	3	2	3	2	3	2	2
9	2	2	2	2	3	3	2	2
10	2	2	3	3	3	2	2	2
11	2	2	2	2	2	2	2	2
12	2	2	2	2	3	3	2	2
13	2	3	3	3	3	3	2	2
14	2	1	2	2	1	1	1	2
15	2	2	2	2	2	3	2	2
16	1	1	3	2	1	1	1	2
17	2	2	2	2	3	3	2	2
18	1	1	2	2	3	2	2	2
19	2	2	2	2	2	2	2	2
20	3	2	3	3	2	3	2	3
21	2	2	2	2	2	3	2	2
22	3	3	2	2	3	2	2	3
23	3	2	3	2	3	2	2	2
24	3	2	2	3	3	2	2	2
25	3	2	3	2	3	3	2	3
26	2	2	3	2	3	3	2	2
27	1	1	1	2	1	1	1	1
28	1	2	1	2	2	1	2	2
29	1	2	2	1	2	3	1	2
30	2	2	3	2	2	3	2	2
31	2	2	3	2	3	2	2	2
32	2	2	2	2	2	3	2	2
33	2	2	3	2	3	3	2	2
34	2	2	2	2	3	2	2	2
35	2	2	3	2	2	2	2	2
36	2	2	2	2	2	3	2	2
37	3	2	3	2	3	3	2	2
38	2	2	2	2	3	2	2	2
39	1	2	1	1	2	2	2	2
40	1	1	2	1	1	1	2	2
41	2	2	2	1	2	2	2	2
42	1	2	2	1	1	2	2	2
43	1	1	1	2	1	1	1	2
44	3	3	2	2	3	2	2	2
45	2	3	3	2	3	2	2	2

Results of Intra-Class Correlation

Reliability can be increased by adding more training to reduce the One-judge Reliability or by adding more judges to reduce the variability of the mean.

Number of Judges needed to reach Aspiration Level of Reliability

Aspiration Level	One Judge Reliability			Number of Judges Needed		
	0.335	0.421	0.399	Mathematics	Reading	Science
0.7	4.6	3.2	3.5	5	4	4
0.8	7.9	5.5	6.0	8	6	7
0.9	17.9	12.4	13.6	18	13	14
0.95	37.7	26.1	28.6	38	27	29

Notes: The minimum number of judges calculation is based on the Spearman Browne Prophecy formula,

$$m = \left\{ \frac{\rho^*}{1-\rho^*} \middle/ \frac{\rho_L}{1-\rho_L} \right\} = \frac{\rho^*(1-\rho_L)}{\rho_L(1-\rho^*)}$$, where ρ^* is the reliability aspired to and ρ_L is the reliability estimate for a single judge.

The two-way analysis assuming both random items and fixed judges gives a result for the mean correlation identical to Cronbach's Alpha, i.e., $\alpha = \frac{\sigma^2_{Bet} - \sigma^2_e}{\sigma^2_{Bet}}$. While SPSS allows the user to select between the random and mixed models, the calculations come out the same with either model. Assuming the judges are fixed would imply these are the only judges that would ever be used so there is no component of variance associated with them. *Random judges* assume the judges used are one of many possible selections of judges; then the variability among judges must be taken into account, which will result in a lower value for the intra-class correlation (or any other measure of reliability.)

For the mixed model (i.e., fixed judges), the intra-class correlation would be calculated identically to Alpha.

$$ICC_{FixedJudges} = \frac{ItemMS - EMS}{ItemMS}$$

For the random model, the correct calculation is:

$$ICC_{RandomJudges} = \frac{ItemMS - EMS}{ItemMS + \frac{(JudgeMS - EMS)}{n}}$$

Calculation Modes

Calculation for two-way model with both questions and judges random

Grade 3:

	Language Usage	
	DF	MS
questions	39	2.19
judges	7	0.67
error	273	0.20
Intra-Class Correlation		.91
Cronbach's Alpha		.91

Grade 4:

	Language Usage	
	DF	MS
questions	39	1.41
judges	7	0.63
error	273	0.19
Intra-Class Correlation		.86
Cronbach's Alpha		.87

Grade 5:

	Language Usage	
	DF	MS
questions	41	1.91
judges	7	0.37
error	287	0.20
Intra-Class Correlation		.90
Cronbach's Alpha		.90

Grade 6:

	Language Usage	
	DF	MS
questions	41	2.23
judges	7	0.58
error	287	0.20
Intra-Class Correlation		.91
Cronbach's Alpha		.91

Grade 7:

	Language Usage	
	DF	MS
questions	44	1.79
judges	7	0.24
error	308	0.18
Intra-Class Correlation		.90
Cronbach's Alpha		.90

Grade 8:

	Language Usage	
	DF	MS
questions	44	1.80
judges	7	0.68
error	308	0.18
Intra-Class Correlation		.89
Cronbach's Alpha		.90

Grade 10:

	Language Usage	
	DF	MS
questions	44	1.08
judges	7	0.90
error	308	0.20
Intra-Class Correlation		.81
Cronbach's Alpha		.82